AD A049250

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TR&A Supplement Report #1 no. 1	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) 1 Oct 76 - 31 Jan 77 ADAPTIVE EXPONENTIAL SMOOTHING.	12	5. TYPE OF REPORT & PERIOD COVERED 10-1-76 to 1-31-77 Supplement to Final Report
		6. PERFORMING ORG. REPORT NUMBER #1
7. AUTHOR(s) Frederick A. Rothe		8. CONTRACT OR GRANT NUMBER(s) ARPA #2518 N00014-76-C-0137
9. PERFORMING ORGANIZATION NAME AND ADDRESS Department of International Relations University of Southern California University Park, Los Angeles, CA. 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS ARPA Order-2518 NR 170/773
11. CONTROLLING OFFICE NAME AND ADDRESS Organizational Effectiveness Research Office of Naval Research (Code 452) 800 No. Quincy St., Arlington, Va. 22217		12. REPORT DATE November 1977
		13. NUMBER OF PAGES 17
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Office of Naval Research Branch Office 1030 East Green Street Pasedena, California 91106	12 19p.	15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

"Approved for public release; distribution unlimited."

DDC
RECEIVED
JAN 11 1978
F

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

"Approved for public release; distribution unlimited."

18. SUPPLEMENTARY NOTES

Sup to A031438 per ONR, code 452, Ms. Duncan

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Adaptive, Early Warning, Monitoring, Forecasting, Prediction, Heuristic, Model.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The adaptive exponential smoothing technique, and its utility for international affairs analysis, is discussed. Both monitoring and forecasting aspects of the technique are described. Calculation procedures are illustrated and "worked through" in an example using international event/interaction data.

AD No.
DDC FILE COPY

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

S/N 0102-014-6601

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

408 614

International Relation Research Institute
School of International Relatons
University of Southern California
November, 1977

ADAPTIVE EXPONENTIAL SMOOTHING
TR&A Supplement Report #1
Threat Recognition and Analysis Project

"Approved for public release; Distribution Unlimited." Reproduction in whole or in part is permitted for any purpose of the United States Government. This research was supported by the Advanced Research Projects Agency of the Department of Defense and was monitored by ONR under Contract No. N00014-76-C-0137. The views and conclusions contained in this document are those of the authors and should not be interpreted as necessarily representing the official policies, either expressed or implied, of the Advanced Research Projects Agency or the U.S. Government.

ADAPTIVE EXPONENTIAL SMOOTHING

Frederick A. Rothe
International Relations Research Institute
University of Southern California
November, 1977

The adaptive exponential smoothing technique currently in use at the International Relations Research Institute is the subject of this writing. Developed originally by Herbert Calhoun (McClelland, et. al., 1971) for the WEIS research endeavor, it is now being refined and tested for early warning research and analysis applications. Not as technical as the Calhoun report, this discussion seeks to present a clear and parsimonious explanation of what adaptive exponential smoothing is and why it is useful for early warning research in international affairs. The reader is referred to the 1971 WEIS report (McClelland, et. al.) for Calhoun's theoretical discussion of the technique and to Richard Beal's (1977) application of it to international crisis analysis. For further theoretical treatment, the basic work by Brown (1962) should be consulted. Wheelwright and Makridakis (1973) give an excellent discussion of its utility in an applied management setting. First is a general discussion of exponential smoothing and its usefulness as a forecasting method using event/interaction data.

At the International Relations Research Institute, exponential smoothing is used to analyze event flows; that is, the pattern and dynamics of event/interaction data. A basic concept orienting this analysis is that the present condition of the event flow is the best indicator of what it will be in the future. The independent variable is the event flow at time T and the dependent variable is the event flow at time T + 1. In conjunction with this concept is the practical need, in monitoring and forecasting, to monitor trends in the recent past in an effort to "look ahead" and project the tendency of the event flow. Again, event flow patterns in recent history are probably the best clues for what they will be in the near future. Time intervals usually range from one week to one month, meaning the predicted value is projected ahead by a small lead time (one to four weeks). Predicted values are made for one to three time intervals ahead. Exponential smoothing seems very appropriate for this type of analysis, where the current characteristics and the historically recent trends in the event flow are most salient for projecting future characteristics and trends.

Exponential smoothing has a central use in modelling techniques also, particularly as used by Forrester and his associates (Forrester, 1961). Although the exponential smoothing technique is the same in systems dynamics, placing most emphasis on the most recent data points and less and less on more historically distant data points, the application is somewhat different. Systems dynamics applications smooth data to be placed in a model making long term projections.

Although the systems dynamics exponential smoothing technique is the same as the smoothing method discussed here, the systems dynamics application of smoothing is not appropriate for event flow analysis.

Exponential smoothing is in the "naive" category of forecasting techniques discussed by Charles Roos (1955). Naive techniques are unsophisticated and are simple projections. Techniques in the naive group are guesses, random methods, trend projections, and autocorrelation. Exponential smoothing is a trend projection technique and clearly belongs in this category. Exponential smoothing is an averaging technique and is similiar in that regard to the well known and frequently used moving average method. Both techniques average a specified number of previous values in the data to make the prediction. The main difference between the two techniques is that exponential smoothing adds weights to the most recent values to make the prediction while the moving average technique does not. The basic idea in the smoothing method is that the most recent value of the data is more critical for making the prediction than historically remote values.

The chief advantage of the moving average technique is its simplicity. The predicted values are the averages of previous observed values. Disadvantages are that it cannot correct for trends or cycles, it does not emphasize the most recent or current observed value, and the entire data sequence must be stored and used in the calculations. Exponential smoothing, on the other hand, can correct for trends and cycles, does give weight to the most recent observed value, and does not require the entire data history to be stored for computations. In addition, although more complex than the moving average technique, it is still easy to use and interpret.

The main assumption inherent in both of these forecasting techniques is that there is some underlying pattern in the values of the variable to be forecast. The values of the variable represent not only the underlying pattern, but random fluctuations as well. Both of these methods seek to distinguish between the random fluctuations or "noise," and the basic underlying pattern in the observations of the variable. By "smoothing" the observations of the variable, extreme fluctuations in the data are eliminated. The forecast is based on the smoothed intermediate values. A forecasting method is succesful if it minimizes the difference between the observed and predicted values. This is to say a succesful forecasting technique projects the pattern in the data fairly closely to what the pattern actually is. By smoothing the data and filtering out the noise, the exponential smoothing technique minimizes the error between the observed and predicted values.

In exponential smoothing, random fluctuations are filtered out of the data by averaging and by adjusting the smoothing coefficient in the exponential smoothing model. By weighting with the coefficient the current value of the data, noise is filtered out of the data and estimates of the underlying process can be made. Most simply, exponential smoothing takes the weighted average of several observed values, and combines that with the weighted, current observed value to give the next predicted value of the variable. Below is the basic exponential smoothing model.

$$Y_{i+1} = (1-a)X_i + aY_i$$

where

a = smoothing coefficient ($0 \le a \le 1$)
Y = the last exponentially weighted average
X = the current observation in the sequence

A hypothetical data set will demonstrate how the technique works. Values are placed in the model to generate the predicted values of the sequence. To get the first projected value, the average of the first three data points is calculated. The data are 43, 28, 35, 37, 22. For this example, the smoothing coefficient is set arbitrarily at .1.

$$X = 43 + 28 + 35 = 106$$
$$X = 35.33$$

Values for a(the smoothing coefficient) are included with the most recent data point to generate the predicted value.

$$Y_{i+1} = .9(37) + .1(35.33)$$
$$Y_{i+1} = 33.3 + 3.53$$
$$Y_{i+1} = 36.83$$

The same process is followed for the next iteration.

$$Y_{i+1} = .9(22) + .1(36.83)$$
$$Y_{i+1} = 19.8 + 3.68$$
$$Y_{i+1} = 23.48$$

The two calculated values(36.83 and 23.48) are the projected values for the fourth and fifth time periods of the data sequence. The first three time periods have no projected values since the first three observed values are used to calculate the initial average used in the fourth time period. Observed and predicted values are displayed below.

43	.
28	.
35	.
37	36.83
22	23.48

Smoothing coefficients are estimated for the data being analyzed. Exponential smoothing models are not generalized for use on any and all data. Each smoothing model is designed to fit the data being analyzed. Brown writes that exponential smoothing starts with "good clean data and a reasonable model to represent the process being forecast. The model is fitted to the data; that is, the coefficients in the model are estimated from the data available to date. So far, the problem is a simple curve-fitting problem. The differences are two. (1) The model should fit current data very well, but it is not important that data obtained a long time ago fit well. (2) The computations are repeated with each new observation" (1962: 88-89).

A major difficulty is estimating the coefficient value to put in the model, since a particular method is not available for finding the coefficient value. A well advised technique is to get to "know" the data by analyzing it with descriptive statistics. Once the pattern and stability of the data are known, the coefficients can be placed in the model, keeping this thought in mind: "If the forecasts are to be stable and are to smooth out random fluctuations, we have shown that one should use a small smoothing coefficient or a large number of observations in the average....On the other hand, when one wants rapid response to a real change in the pattern of the observations, then a large value of the smoothing constant is appropriate" (Wheelwright and Makridakis, 1973: 118). The best way to find the appropriate coefficient value is on the anvil of experiment. Different values of the coefficient should be tried until the one is found that minimizes the difference between the observed and predicted values.

Up to this point, this discussion has been concerned with a single model that fits relatively stable data and is first order. In the above demonstration, a single model using first order exponential smoothing was used to calculate the predicted values. First order means the predicted values are derived from the observed values. A problem with first order exponential smoothing, whatever the smoothing coefficient, is the "smoothed averages will generally lag behind a steadily rising or falling trend, resulting in a cumulative error in prediction" (Calhoun in McClelland, et. al., 1971: 278). Second order smoothing is used to correct this error by simply smoothing over the first smoothed averages using the same weighting scheme. This trend correction factor is the difference between first and second order smoothing. The second order average is found by smoothing the first order, as follows:

$$Z_{i+1} = aY_i + (1-a)Z_{i-1}$$

where

a = smoothing coefficient ($0 <= a <= 1$)
Y = the last exponentially smoothed average
Z = the second order average

Second order smoothing is demonstrated below, using the same hypothetical data as before, and their predicted first order values.

ORIGINAL OBSERVED	PREDICTED FIRST ORDER
43	35.33
28	35.33
35	35.33
37	36.83
22	23.48

The first iteration is calculated below.

$$Zi+1 = .9(36.83) + .1(35.33)$$
$$Zi+1 = 33.74 + 3.53$$
$$Zi+1 = 36.67$$

Note that 35.33 is the average from the first three observed values. In the first iteration, it is used as the last exponentially smoothed average. The first order average is used in place of the second order average on the first iteration also. Iterations following the first are calculated using the second order average and the last exponentially smoothed average. The next iteration is calculated below.

$$Zi+1 = .9(23.48) + .1(36.67)$$
$$Zi+1 = 21.13 + 3.66$$
$$Zi+1 = 24.79$$

Second order observed and predicted values follow.

43	.
28	.
35	.
37	36.67
22	24.79

In this hypothetical example, the second order predicted values are not as close as the first order predicted values. The reason is that second order smoothes the first order predictions, causing second order predictions not to fit the observed values as closely. Second order smoothing minimizes the cumulation error in first order, caused by steadily rising or falling trends, however.

Third order exponential smoothing can be used also. Third order is applicable when no steadily rising or falling trends are observed in the data. Third order projected values are not as close to the observed values as the other orders, because third order smoothes the second order projected values. Third order smoothing is calculated the same as the other two, except that it uses the last exponentially smoothed second order value and the third order average (after the first iteration).

In many ways, first order exponential smoothing using one model may not be suitable for data used in international affairs research. First, the data may change substantially over time, thus making the initial smoothing model obselete, and second, many of the changes or sudden perturbations in the data may be substantive and should not be discounted as "noise" or random fluctuations. Calhoun recognized this (1971: 285-291) and developed four models that are designed to fit different variations in event interaction data. Calhoun developed these models in response to the basic premise of most prediction models. Prediction models are designed to fit two components of the data. One component is the underlying pattern, or permanent component, and the other is the noise, or random component. Departures from one of these components in the data, in the form of abrupt shifts or discontinuities, usually destroy the reliability of the predictive model. Discontinuities and abrupt changes in event/interaction data are substantive and should be incorporated into the predictive models. Through analysis of the discontinuities in the WEIS data, Calhoun discovered four separate components of the data. Each component of the data is treated with a separate exponential smoothing model for its unique qualities. Therefore, when a discontinuity in the data is encountered, it is treated not as either the permanent or random component, but as a separate substantive component. Another insight of the notion of different but not random components in the data is that much of the randomness in international relations data "appears to be an integral part of the phenomena. Attempts to compensate for randomness or to de-randomize are tantamount to throwing away the most essential parts of the data" (Calhoun in McClelland, et. al., 1971:285-286). Other models may be necessary to fit the many variations of the data, or the models themselves may have to be adjusted. Admittedly experimental, the models are:

(1) The moderate up-down fluctuation model: moderate volume which is relatively autocorrelated.

(2) The random drastic upturn-downturn model: moderate to low volume with isolated peaks of three standard deviations or greater.

(3) The gradual monotone upward or downward model: moderate to high volume with a definite trend component.

(4) The drastic up-down model: semi-autocorrelated with peaks and troughs of two standard deviations or more. (Calhoun in McClelland, et.al., 1971: 285).

Model I is one that occurs frequently in the WEIS data. It describes basically the routine, ongoing, maintenance activity in the system, a dyad, or groups of nations. Few disturbances occur in the data of this model. One standard deviation is the operational threshold value for the model. If the standard deviation is exceeded at any point in time, then another of the smoothing models is invoked. Model I is defined operationally as the model whose smoothing parameters, X and Y (from the earlier part of this writing), remain within a range of one standard deviation from one point in time to the next.

Model II is difficult to predict. Patterns of data suitable for this model are infrequent but usually signal important occurrences. It is designed to capture departures from the routine activity of the data. It is based on the difference between the previous observed and predicted values. These differences are determined by standard deviations also, with the difference between the observed values being greater than three standard deviations and the difference between the predicted and observed values being less than two standard deviations.

Model III is much like Model II but requires the data to have a definite upward or downward trend. If a trend has occurred over the last four or fewer time periods, the latest three time periods are detrended and the smoothing equation is similar to Model II. If no trend exists, then Model II is invoked.

Model IV finally, is a severe case of Model I. Threshold value size is the chief difference between the two models. Model IV tests to see if smoothing parameter X, the difference between observed values, is greater than three standard deviations and parameter Y, the difference between observed and predicted values, is greater than two standard deviations.

Each data point is tested against these models. If the conditions for a model are met, then its appropriate smoothing equations are invoked. If the conditions for one model are not met, the conditions for the next are tested. This procedure continues until one model is invoked and the prediction is made. In this fashion, the basic exponential smoothing technique adapts to changes in the data, and invokes the smoothing model that fits the characteristics of the data.

An example of the output from the adaptive exponential smoothing program in use at the International Relations Research Institute is presented shortly, but first the operational forms of the four models are given.

MODEL	CONDITIONS
I. Moderate Up-Down	$\lvert F\rvert <= 1a$ and $\lvert P\rvert <= 1a$
II. Isolated Drastic Upturn or Downturn	$\lvert F\rvert => 3a$ and $\lvert P\rvert <= 2a$
III. Gradual Monotone Up and Down	($\lvert F\rvert => 1a$ or $\lvert F\rvert <= 3a$) and ($\lvert P\rvert => 1a$ or $\lvert P\rvert <= 2a$)
IV. Drastic Up-Down	$\lvert F\rvert > 3a$ or $\lvert P\rvert > 2.5a$

where F = (Xi - Xi-1)
P = (Yi - Xi)
a = standard deviation
X = observed value
Y = predicted value

Below is the output taken from a PLI program written by the author and Richard S. Beal. It is a descendant of Calhoun's PLI/CPS interactive program and Beal's BASIC language program. The data are from Professor Charles McClelland's EFI33 Index. The Event Flow Indicator (EFI), version 33, was developed by McClelland to be used for monitoring and early warning. It is a measure based on the volume and variety of activity in the international system. After considerable experimentation McClelland has settled on using eight of the major event(COMBEVENT) categories, out of the possible 22 in the WEIS collection, for the EFI calculations. McClelland(1976) has demonstrated that when the EFI score drops below 500, there is a dangerous situation in the international setting. Scores from approximately 530 to 500 are interpreted as possible precursors or early warning of dangerous situations. The observed values in the example are EFI33 scores, calculated from The TIMES of London data on a weekly basis, beginning on January 2, and ending on October 1, of 1977.

DATE	TIME	MODEL	OBSERVED	ESTIMATE	RESIDUAL	%RESIDUAL
770108	1	0	686.80	0.00	0.00	0.00
770115	2	0	684.10	0.00	0.00	0.00
770122	3	0	696.30	0.00	0.00	0.00
770129	4	2	614.30	700.32	-86.02	-14.00
770205	5	3	586.20	568.31	17.89	3.05
770212	6	1	559.20	575.49	-16.29	-2.91
770219	7	1	560.60	550.67	9.93	1.77
770226	8	2	604.70	476.37	128.33	21.22
770305	9	3	566.70	543.09	23.61	4.17
770312	10	2	643.20	498.14	145.06	22.55
770319	11	3	686.20	621.24	64.96	9.47
770326	12	3	596.60	714.32	-117.72	-19.73
770402	13	3	643.80	581.74	62.06	9.64
770409	14	3	579.00	637.30	-58.30	-10.07
770416	15	3	602.50	541.83	60.67	10.07
770423	16	3	620.80	567.13	53.67	8.65
770430	17	3	573.60	604.45	-30.85	-5.38
770507	18	2	675.00	535.41	139.59	20.68
770514	19	3	641.90	692.23	-50.33	-7.84
770521	20	3	520.70	663.57	-142.87	-27.44
770528	21	3	579.80	459,29	120.51	20.79
770604	22	3	589.70	526.15	63.55	10.78
770611	23	3	552.00	558.54	-6.54	-1.19
770618	24	3	582.60	505.47	77.13	13.24
770625	25	3	611.90	550.43	61.47	10.05
770702	26	3	626.20	608.75	17.45	2.79
770709	27	3	653.80	642.30	11.50	1.76
770716	28	1	659.00	656.97	2.03	0.31
770723	29	2	528.00	698.78	-170.78	-32.35
770730	30	3	535.80	488.95	46.85	8.74
770806	31	3	597.80	469.29	128.51	21.50
770813	32	3	564.10	572.92	-8.82	-1.56
770820	33	1	558.50	561.98	-3.48	-0.62
770827	34	1	552.60	555.60	-3.00	-0.54
770903	35	2	593.70	521.74	71.96	12.12
770910	36	3	628.50	585.68	42.82	6.81
770917	37	3	603.50	654.02	-50.52	-8.37
770926	38	3	643.50	622.69	20.81	3.23
771001	39	2	512.80	680.76	-167.96	-32.75
771008	40	3	0.00	474.43	0.00	0.00
771015	41	3	0.00	716.21	0.00	0.00
771022	42	3	0.00	418.25	0.00	0.00

No models are invoked and no estimates are made for the first three time periods. As in the above example, values for the first three time periods are used to calculate the initial average for the predicted (ESTIMATE) value in time period 4. Observed values for time periods 2, 3, and 4 are used to make the prediction in time period 5 and so on.

Calhoun realized that exponential smoothing, as a curve-fitting technique, did not make projections ahead. This means that after the initial average was calculated from the first three observed values, the fourth observed value was used to calculate the first predicted value, also in the fourth time period. Calhoun altered this so the third observed value is used, after the calculation of the average, to calculate the first predicted value in the fourth time period. Calhoun's smoothing technique is designed to fit the pattern of the data and to project the next value. Therefore, in time period 40, the predicted value(ESTIMATE) is calculated from the observed value in time period 39. Time periods 40, 41, and 42 contain no observed values because the data ended in time period 39. Predicted values are still made by assigning the ESTIMATE value from time period 40 to OBSERVED time period 40. This is used to calculate the ESTIMATE in time period 41. The same procedure is followed for the desired amount of predictions.

Model 2 is the model invoked in time period 4. It's conditions are calculated in the following fashion.

770101	686.8
770115	684.1
770122	696.3

MEAN(M) = 689.06
STANDARD DEVIATION(SD) = 5.21
F = 614.30 - 696.30 = 82.00
P = 689.06 - 696.3 = 7.24

Both the mean and the standard deviation are calculated in moving groups of three. From the earlier discussion, F is the difference between OBSERVED(T) and OBSERVED(T-1), and P is the difference between ESTIMATE(T-1) and OBSERVED(T-1). The program calculates these in "absolute" differences, thus eliminating all negative values.

Model 3 is invoked in time period 5 and the calculations of its parameters are displayed below.

770115	684.1
770122	696.3
770129	614.3

M = 644.9
SD = 36.13
F = 586.20 - 614.30 = 28.10
P = 700.32 - 614.30 = 86.02

Finally, calculations for parameters in time period 6 and Model 1 follow.

770122	696.3
770129	614.3
770205	586.2

M = 632.27
SD = 46.71
F = 559.20 - 586.20 = 27.00
P = 575.49 - 559.20 = 17.89

Calhoun's smoothing equations are developed, through testing simulated data and intuition, to fit the unique qualities of event/interaction data. The different qualities are represented by the four models. The equations used in the PLI program are not exactly like those discussed at the outset of this paper, but are designed to make the best fit(minimize error) for each of the models. Calhoun experimented with and tested the models until he found these four, which represent the four different components of the data. Calhoun developed his smoothing equations based on the work of others (Calhoun in McClelland, et. al., 1971: 295). He tailored his equations with a correction factor that adds or subtracts a predetermined amount from the previously smoothed average. Correction of the data was suggested by the constant movement in the values of the data over successive time periods. The smoothing equations are presented here in their operational form. The main smoothing equation for first order is given first, followed by its extension for each of the models.

```
FNA = (.9 * OBSERVED(I-1)+(.1 * ESTIMATE(I-1))
FNB = (.1 * FNA ) + (.9 * NEWMEAN)

MODEL I
     ESTIMATE(I) = FNA + (.1 * (FNA - FNB))

MODEL II
     ESTIMATE(I) = FNA + (.9 * (FNA - FNB))

MODEL III
     ESTIMATE(I) = FNA + (.9 * (FNA - FNB))

MODEL IV
     ESTIMATE(I) = FNA + (.3 * (FNA - FNB))
```

FNA and FNB are variables used to hold the results of the equations, and I is the time period. For example, in the first iteration I is equal to 4 because the prediction is for time period 4. NEWMEAN is the new moving average calculated for each new group of three time periods. The first two equations are computed for every iteration through the data. They are then placed in the equation for each model and the predicted value comes from these equations. Important here is the calculation for second order smoothing. FNB is substituted for NEWMEAN, thus smoothing over the first average. The first two equations then look like this.

```
FNA = (.9 * OBSERVED(I-1)+(.1 * ESTIMATE(I-1))
FNB = (.1 * FNA) + (.9 * FNB)
```

The first equation, FNA, is the regular smoothing equation, identical to the one displayed in the beginning of the paper. Equation FNB was developed by Calhoun to further smooth the data and was derived from his experimentation with the WEIS data. This basic smoothing procedure is tailor made for the WEIS collection, thus fulfilling the major requirement of exponential smoothing to fit the data being analyzed.

Examples of the calculations for time periods 4, 5, and 6 of the example follow. These calculations were done by hand with the aid of a hand calculator, and they are not as accurate to the last decimal as the calculations done by the computer.

TIME	MODEL	OBSERVED	ESTIMATE
1		686.80	689.07
2		684.10	689.07
3		696.30	689.07
4	2	614.30	
5	3	586.20	
6	1	559.20	

Model numbers are from the earlier calculations and the ESTIMATE values are the mean for the first three time periods. FNB is calculated twice in the first iteration because this is second order. After that, FNB is placed in the calculations from the previous time period.

First iteration for Time Period 4:

FNA = .9(696.3) + .1(689.07)
= 626.67 + 68.907
<u>695.577</u>

FNB = .1(695.577) + .9(689.07)
= 69.577 + 620.163
= 689.720

FNB is calculated again on the first iteration.

FNB = .1(695.577) + .9(689.720)
= 69.577 + 620.748
<u>690.325</u>

Model II is invoked in time period 4.

ESTIMATE = 695.577 + .9(695.577 - 690.325)
= 695.577 + .9(5.251)
= 695.577 + 4.726
<u>700.30</u>

Second iteration for Time Period 5:

FNA = .9(614.3) + .1(700.30)
= 522.87 + 70.03
<u>622.90</u>
FNB = .1(622.90) + .9(690.325)
= 62.29 + 621.293
<u>683.583</u>

Model III is invoked for time period 5.

```
ESTIMATE = 622.90 + .9(622.90 - 683.583)
         = 622.90 + .9(-60.683)
         = 622.90 - 54.614
           568.28
```

Third iteration for Time Period 6:

```
FNA = .9(586.2) + .1(568.28)
    = 527.58 + 56.828
      584.408
FNB = .1(584.408) + .9(683.583)
    = 58.440 + 615.224
      673.665
```

Model I is invoked for time period 6.

```
ESTIMATE = 584.408 + .1(584.408 - 673.665)
         = 584.408 + .1(-89.257)
         = 584.408 - 8.925
           575.48
```

It bears repeating that the OBSERVED value used in each calculation is from the previous time period. On the last iteration, the OBSERVED is 696.3, from the third time period, used for calculating the ESTIMATE value in the fourth time period. The final underlined figure for each example is the predicted (ESTIMATE) value for each of the time periods. Except for the last decimal places, these values are identical to those calculated by the computer in the above example.

Residuals are used to analyze how closely the predicted values fit the observed values. In time period four, the residual is -86.02, but it is difficult to interpret what the -86.02 means unless it can be compared to other residuals. For this reason, the %RESIDUAL value was added. %RESIDUAL is calculated by dividing the residual by the observed value and multiplying the quotient by 100. Time period 4 is calculated below as an example.

```
%RESIDUAL = (RESIDUAL/OBSERVED) * 100
%RESIDUAL = (-86.02/614.3) * 100
%RESIDUAL = -.14 * 100
%RESIDUAL = -14.00
```

%RESIDUAL gives the percent difference between the observed and predicted values. While -86.02 appears to be a large "miss" for time period 4, the %RESIDUAL indicates the ESTIMATE value "missed" the OBSERVED value by only -14 percent of the observed value. Notice the largest %RESIDUALS (-32.25 and -32.75) are in time periods 29 and 39. The ESTIMATE values for these time periods can be interpreted as "fitting" 133% of the OBSERVED value. It seems then that this technique is fairly "successful." In this example at least, the ESTIMATE values predicted to approximately plus or minus 30% of the observed values at the very worst. Accuracy of this type may not always hold, and further testing of the technique is necessary, but this first example with "real" data does serve as a source of encouragement.

Confidence in the technique is enhanced even more by the calculation of another percentage. As each %RESIDUAL is calculated, it is checked to see if it is equal to or less than absolute 20.00. The time periods for this occurrence are tabulated. The total is divided by the number of time periods where a %RESIDUAL is calculated. In this example, there are 36 of 42 weeks, excluding the first and last three, with a %RESIDUAL, and of that total, 28 weeks had %RESIDUALs of 20.00 or less. This means that, for this example, the ESTIMATE values predicted, accounted for, or fitted, plus or minus 20 percent of the OBSERVED values 77 percent of the time. For this data, if the general pattern continues, the analyst can have 77% "confidence," or be 77% "sure" that the next observed value has been predicted within 20% over or under its actual value by the ESTIMATE value.

Application of this "confidence" value may be of particular use for forecasting several time periods ahead. Time periods 40, 41, and 42 seem to fluctuate quite wildy, due perhaps to the calculation procedure. Experience with the data indicates that such fluctuations are extremely uncharacteristic and that the predicted values should be tempered. Since 80% to 120% of the OBSERVED values are predicted 77% of the time, it may be advisable to add 20% of the ESTIMATE to the ESTIMATE if it is low or subtract 20% of the ESTIMATE from the ESTIMATE if it is high. Unrefined and intuitive, this measure produces ESTIMATE values of 569.31 for time period 40, 592.97 for 41, and 501.89 for the final time period, 42. These values appear more in line with the pattern of the data.

In regression analysis, the standard error of estimate is used to determine prediction accuracy. As is well known, it may be interpreted as an average residual, or the average error in the regression analysis. A standard error of estimate is calculated for the exponential smoothing technique for the same purpose. In exponential smoothing, the standard error of estimate is calculated using only those time periods that have both observed and predicted values. In the above example, both the first three time periods and the last three time periods are excluded from the computation. The first three are excluded because they have no predicted values and the last three are not included since they have no observed values. In this example, the standard error of estimate is 81.03.

Not only is this a confidence measure used to see how closely the predicted values fit the observed values, but it is a good measure to use for refining the smoothing technique. As different smoothing coefficients are tried, and different numbers of observations are used for the moving average, this measure can be compared over the different trials until the best fit is found.

The standard error of estimate can also be used to adjust the values forecasted several time periods ahead. It is certainly more rigorous than the method described previously and can be used in the same fashion; that is, add the standard error of estimate to low predicted values and subtract it from high predicted values. This gives values of 555.46 for time period 40, 635.18 for time period 41, and 499.28 for the last time period.

An additional aspect of the adaptive exponential smoothing technique is its heuristic quality. Not mentioned earlier, it is exercised under the MODEL heading in the example. Observing, tabulating, and analyzing the pattern and frequency of the different models may indicate sequences of action. Each of the models is invoked under specific conditions, so the model numbers themselves indicate rough characteristics of the data. Repetition of the sequences, if any are found, particularly before and after known historical occurrences, may lead to classification schemes for the sequences and investigation of the factors contributing to the sequences. Big power involvement, regional disturbances, total number of actors acting, total event volume, total number of targets receiving, different events enacted in various orders, and changes, differences, and ratios among some or all these factors may contribute to understanding how troublesome, dangerous situations, requiring early warning, emerge and dissipate.

To begin the summary of this paper, the difference between regular and adaptive exponential smoothing is stated once again. Regular exponential smoothing uses the same exponential model for the entire history of the data. Adaptive exponential smoothing, however, invokes a different model, one of four in the technique described here, for each value in the data sequence.

The strengths of second order adaptive exponential smoothing in international relations research are these. First, and not mentioned previously, exponential smoothing does not require "rich" data that other longitudinal techniques, like time-series analysis(e.g., Hill in McClelland, et.al., 1971), require to be useful. Exponential smoothing techniques can be used on any data collected at the same intervals over time. Given the difficulties with getting "good" data in international affairs, this aspect is appealing.

Second, exponential smoothing is easy to use and interpret. It is necessary only to fit the smoothing model to the data. This is done in an experimental fashion until the model is succesful. A succesful model reduces error as much as possible between the observed and predicted values.

Finally, adaptive exponential smoothing accounts for perturbations in the data and does not discount them as noise.

Adaptive smoothing attempts to capture the sudden, and substantive changes in the data by using different smoothing models. This requires the analyst to know the data and to develop models according to the demands of the data. The four models discussed briefly here have proven fairly successful, but they may be adjusted or eliminated as more successful models are developed.

Because of these properties, the technique seems well suited for early warning research. It can be used for analyzing a dyad's activity or it can be used for analyzing total volume in the system. In addition to prediction, the adaptive technique can be used for monitoring by observing which models are invoked over time and in what order. Periods of time with several different models may be studied more closely to determine what circumstances are present in the system, causing the various models to be invoked.

Adapting to fluctuations in the data seems to be the most appealing aspect of this type of smoothing for early warning research. But forecasting the dynamic of international affairs is not without considerable difficulty. If analysts of international affairs knew why the affairs moved through time as they do, then the forecasting difficulty would be considerably diminished. Until that knowledge is developed, however, accumulating understanding on how international affairs work is essential. The heuristic qualities of the technique discussed here lend themselves to that task.

References

Beal, Richard Smith 1977
Systems Analysis of International Crises: Event Analysis of Nine Pre-Crisis Threat Situations, 1948-1962.
Unpublished Ph.D. Dissertation. University of Southern California, Los Angeles, California.

Brown, Robert G. 1962
Smoothing, Forecasting, and Predictions of Discrete Time Series.
Prentice-Hall: Englewood Cliffs, New Jersey.

Forrester, Jay W. 1961
Industrial Dynamics.
M.I.T. Press: Cambridge, Mass. (See Appendix E.)

McClelland, Charles A. 1976
"Warnings in the International Event Flow: EFI and ROZ as Threat Indicators." Los Angeles: University of Southern California, (mimeo).

McClelland, Charles A., and others, 1971
The Management and Analysis of International Event Data: A Computerized System for Monitoring and Projecting Event Flows.
WEIS Technical Report, University of Southern California: Los Angeles, California.

Roos, Charles 1955
"Survey of Economic Forecasting Techniques,"
Econometrica
(October): 363-395.

Wheelwright, Steven C. and Spyros Makridakis 1973
Forecasting Methods for Management.
John Wiley and Sons, New York.